Prototypical Cross-domain Self-supervised Learning for Few-shot Unsupervised Domain Adaptation

Xiangyu Yue[1*], Zhangwei Zheng[2*], Shanghang Zhang[1], Yang Gao[3], Trevor Darrell[1], Kurt Keutzer[1], Alberto Sangiovanni-Vincentelli[1]

[1]University of California, Berkeley [2]Nanjing University [3]Tsinghua University

CVPR VIRTUAL JUNE 19-25

Motivation

Previous: Unsupervised Domain Adaptation

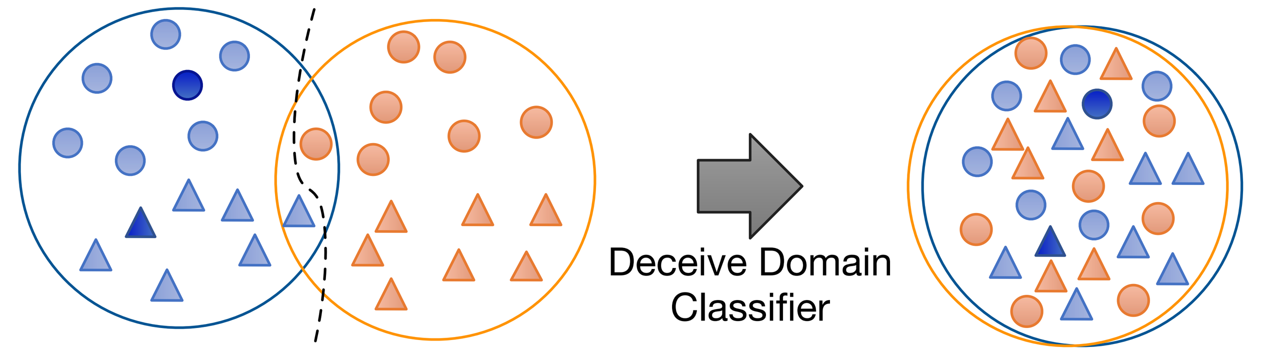
- **Fully-labeled** Source: *sometimes it is hard even to get labels in the source domain, e.g. medical Imaging and pixel-level annotations*
- Unlabeled Target

Few-shot Unsupervised Domain Adaptation

- Source with **few-shot labels**
- Unlabeled Target

Overall Comparison

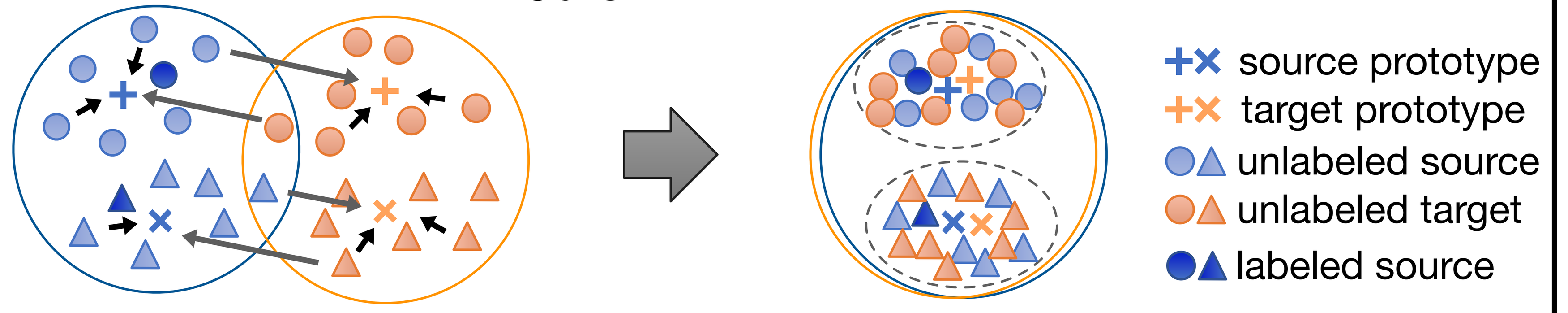
Previous method



Train domain classifier   Update features

With few-shot labels in the source, it is hard to learn a good classifier in both source and target.

Ours

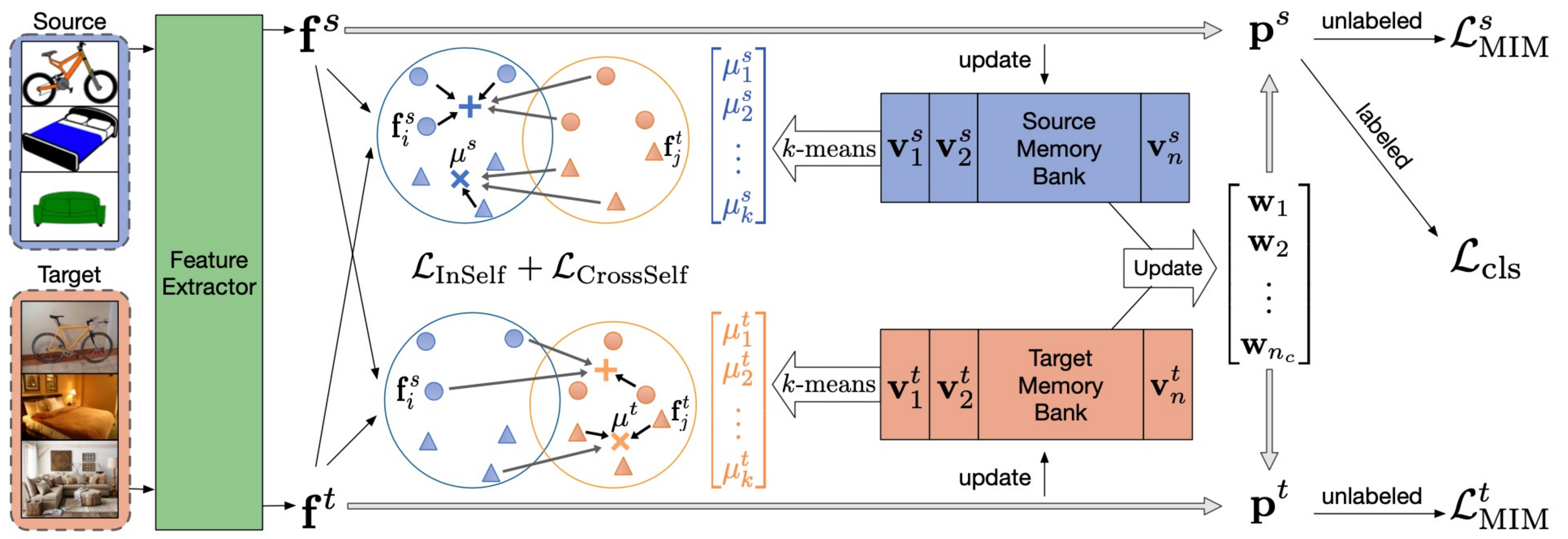


PCS   After adaptation

Our method perform prototypical SSL both in-domain and cross-domain and learn better classifier in both domains.

Method

PCS Framework



In-Domain SSL

$$P^s_{i,j} = \frac{\exp(\mu^s_j \cdot \mathbf{f}^s_i/\phi)}{\sum_{r=1}^{k}\exp(\mu^s_r \cdot \mathbf{f}^s_i/\phi)} \qquad \mathcal{L}_{\text{PC}} = \sum_{i=1}^{N_s+N_{su}} \mathcal{L}_{CE}(P^s_i, c_s(i)) + \sum_{i=1}^{N_{tu}} \mathcal{L}_{CE}(P^t_i, c_t(i))$$

Cross-Domain SSL

$$P^{s\to t}_{i,j} = \frac{\exp(\mu^t_j \cdot \mathbf{f}^s_i/\tau)}{\sum_{r=1}^{k}\exp(\mu^t_r \cdot \mathbf{f}^s_i/\tau)} \qquad \mathcal{L}_{\text{CrossSelf}} = \sum_{i=1}^{N_s+N_{su}} H(P^{s\to t}_i) + \sum_{i=1}^{N_{tu}} H(P^{t\to s}_i)$$

Adaptive Prototypical Classifier Learning

$$\hat{\mathbf{w}}^s_i = \frac{1}{|\mathcal{D}^{(i)}_{s+}|}\sum_{\mathbf{x}\in\mathcal{D}^{(i)}_{s+}} m(\mathbf{x}); \quad \hat{\mathbf{w}}^t_i = \frac{1}{|\mathcal{D}^{(i)}_{tu}|}\sum_{\mathbf{x}\in\mathcal{D}^{(i)}_{tu}} m(\mathbf{x}) \qquad \mathbf{w}_i = \begin{cases} unit(\hat{\mathbf{w}}^s_i) & \text{if } |\mathcal{D}^{(i)}_{tu}| < t_w \\ unit(\hat{\mathbf{w}}^t_i) & \text{otherwise} \end{cases}$$

Results

Method	Office: Target Acc. on 1-shot / 3-shots						
	A→D	A→W	D→A	D→W	W→A	W→D	Avg
SO	27.5 / 49.2	28.7 / 46.3	40.9 / 55.3	65.2 / 85.5	41.1 / 53.8	62.0 / 86.1	44.2 / 62.7
MME [57]	21.5 / 51.0	12.2 / 54.6	23.1 / 60.2	60.9 / 89.7	14.0 / 52.3	62.4 / 91.4	32.3 / 66.5
CDAN [42]	11.2 / 43.7	6.2 / 50.1	9.1 / 65.1	54.8 / 91.6	10.4 / 57.0	41.6 / 89.8	22.2 / 66.2
CAN [35]	25.3 / 48.6	26.4 / 45.3	23.9 / 41.2	69.4 / 78.2	21.2 / 39.3	67.3 / 82.3	38.9 / 55.8
MDDIA [32]	45.0 / 62.9	54.5 / 65.4	55.6 / 67.9	84.4 / 93.3	53.4 / 70.3	79.5 / 93.2	62.1 / 75.5
CDS [36]	33.3 / 57.0	35.2 / 58.6	52.0 / 67.6	59.0 / 86.0	46.5 / 65.7	57.4 / 81.3	47.2 / 69.3
DANN + ENT [17]	32.5 / 57.6	37.2 / 54.1	36.9 / 54.1	70.1 / 87.4	43.0 / 51.4	58.8 / 89.4	46.4 / 65.7
MME + ENT	37.6 / 69.5	42.5 / 68.3	48.6 / 66.7	73.5 / 89.8	47.2 / 63.2	62.4 / 95.4	52.0 / 74.1
CDAN + ENT	31.5 / 68.3	26.4 / 71.8	39.1 / 57.3	70.4 / 88.2	37.5 / 61.5	61.9 / 93.8	44.5 / 73.5
CDS + ENT	40.4 / 61.2	44.7 / 66.7	66.4 / 73.1	71.6 / 90.6	58.6 / 71.8	69.3 / 86.1	58.5 / 74.9
CDS + MME + ENT	39.4 / 61.6	43.6 / 66.3	66.0 / 74.5	75.7 / 92.1	53.1 / 73.0	70.9 / 90.6	58.5 / 76.3
CDS / MME + ENT†	55.4 / 75.7	57.2 / 77.2	62.8 / 69.7	84.9 / 92.1	62.6 / 69.9	77.7 / 95.4	65.3 / 80.0
CDS / CDAN + ENT†	53.8 / 78.1	65.6 / 79.8	59.5 / 70.7	83.0 / 93.2	57.4 / 64.5	77.1 / 97.4	66.1 / 80.6
PCS (Ours)	**60.2 / 78.2**	**69.8 / 82.9**	**76.1 / 76.4**	**90.6 / 94.1**	**71.2 / 76.3**	**91.8** / 96.0	**76.6 / 84.0**
Improvement	+4.8 / +0.1	+4.2 / +3.1	+9.7 / +1.9	+5.7 / +0.9	+8.6 / +6.4	+14.1 / -1.4	+10.5 / +3.4

Method	VisDA: Target Acc. (%)	
	0.1% Labeled	1% Labeled
SO	47.9	51.4
MME [57]	55.6	69.4
CDAN [42]	58.0	61.5
MDDIA [32]	68.9	71.3
CAN [35]	51.3	57.2
CDS [36]	34.2	67.5
DANN + ENT [17]	44.5	50.2
MME + ENT	54.0	66.1
CDAN + ENT	57.7	58.1
CDS + ENT	49.8	75.3
CDS + ENT + MME	60.0	78.3
CDS / MME + ENT†	62.5	69.4
CDS / CDAN + ENT†	69.0	69.1
PCS (Ours)	**78.0**	**79.0**
Improvement	+9.0	+0.7

Method	DomainNet: Target Acc. (%)							
	R→C	R→P	R→S	P→C	P→R	C→S	S→P	Avg
1-shot labeled source								
SO	18.4	30.6	16.7	16.2	28.9	12.7	10.5	19.1
MME [57]	13.8	29.2	9.7	16.0	26.0	13.4	14.4	17.5
CDAN [42]	16.0	25.7	12.9	12.6	19.5	7.2	8.0	14.6
MDDIA [32]	18.0	30.6	15.9	15.4	27.4	9.3	10.2	18.1
CAN [35]	18.3	22.1	16.7	13.2	23.9	11.1	12.1	16.8
CDS [36]	16.7	24.4	11.1	14.1	15.9	13.4	19.0	16.4
CDS + ENT	21.7	30.1	18.2	17.4	20.5	18.6	22.7	21.5
CDS + MME + ENT	21.2	28.8	15.5	15.8	17.6	19.0	20.7	19.8
PCS (Ours)	**39.0**	**51.7**	**39.8**	**26.4**	**38.8**	**23.7**	**23.6**	**34.7**
Improvement	+17.3	+21.1	+21.6	+9.0	+9.9	+4.7	+0.9	+13.2
3-shots labeled source								
SO	30.2	44.2	25.7	24.6	49.8	24.2	23.2	31.7
MME [57]	22.8	46.5	14.5	25.1	50.0	20.1	24.9	29.1
CDAN [42]	30.0	40.1	21.7	21.4	40.8	17.1	19.7	27.3
MDDIA [32]	41.4	50.7	37.4	31.4	52.9	23.1	24.1	37.3
CAN [35]	28.1	33.5	25	24.7	46.9	23.3	20.1	28.8
CDS [36]	35.0	43.8	36.7	34.1	36.8	31.1	34.5	36.0
CDS + ENT	44.5	52.2	40.9	40.0	47.2	33.0	40.1	42.5
CDS + MME + ENT	43.8	54.9	41.1	38.9	45.9	32.8	38.7	42.3
PCS (Ours)	**45.2**	**59.1**	**41.9**	**41.0**	**66.6**	31.9	37.4	**46.1**
Improvement	+0.7	+6.9	+0.8	+1.0	+13.7	-0.9	-2.7	+3.6

Summary

We propose a novel Prototypical Cross-domain Self-supervised learning framework (**PCS**) for few-shot unsupervised Domain Adaptation, setting a new **State of the Art** for FUDA.

← Project Link